文章编号：2096-1618(2024)02-0178-05

基于网络安全态势感知的四川气象信息网络检测防御技术研究

田 娟， 方国强， 何星庭， 谢银海
(四川省气象探测数据中心，四川 成都 610072)

摘要：基于大数据、机器学习等技术，构建四川气象信息网络安全态势感知系统。系统能全局感知网络状态，提高了四川省气象部门网络监控、应急响应能力，为合理决策提供支持。介绍了网络安全态势感知技术，并阐述气象信息网络及其安全防御现状和存在风险，从实践角度出发介绍该系统及其关键技术，并对系统应用进行总结和展望。

关 键 词：网络安全；网络安全态势感知；气象信息网络；检测防御
中图分类号：TP309.5 **文献标志码**：A
doi：10.16836/j.cnki.jcuit.2024.02.008

0 引言

四川气象信息网络系统是集办公、科研、数据分析在内的多种应用系统，是全省气象预测、预报、服务的基础支撑。在网络安全新形势下，面对日益多样化、复杂化的网络攻击，传统防御工具构成的安全防御体系已到达瓶颈。网络安全态势感知[1](network security situation awareness，NSSA)是网络安全领域的研究热点，能全天候、全方位对网络进行监视，并评估网络安全态势，为网络管理提供决策依据。将该技术应用于四川气象信息网络，构建四川气象信息网络安全态势感知系统，提高了网络监控能力、安全事件响应能力和网络安全发展趋势预测能力。

1 网络安全态势感知技术

网络态势感知技术[2]于 1999 年由 Bass 首次提出，但其概念却始终得不到统一。本文引用石乐义等[3]对网络安全态势感知的定义，认为它能对网络安全要素进行综合分析，评估网络安全状况并预测其发展趋势，是保障网络安全的有效手段。

目前，国内外大多数网络安全态势感知研究以传统 3 层模型[4]为基础，从不同角度进行了丰富细化。3 层模型框架如图 1 所示。

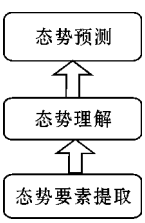

图 1 网络安全态势感知 3 层模型

态势要素提取层：通过各种工具，采集获取影响系统安全的要素，为后续工作提供数据支撑。

态势理解层：基于下层提取要素，采用分类、关联分析等技术对数据进行处理融合，分析安全事件之间的相关性，全局性评估网络安全态势值。

态势预测层：根据当前网络环境状态，结合网络安全的历史数据，预测未来一段时间网络安全发展趋势。

2 四川气象信息网络系统

四川气象信息网络系统是以四川省气象探测数据中心为主中心、市(州)气象局为分中心、县局(站)为终端的三级网络架构。系统实现了中国电信、中国移动双运营商专线加 VPN 应急备份链路，保障各种气象数据的传输。同时，四川省气象局是全国气象系统内拥有市(州)级和县级台站数量最多的省份，站网规模在全国省级气象信息广域网系统中位居第一。

在安全防御方面，系统采用传统方式进行网络安全防御。包括但不限于增设网络安全设备——漏洞扫描、入侵检测、防火墙等；将互联网出口统一到省局，禁止市县互联网出口，减少系统遭受来自互联网的攻击；购买技术厂家终端安全防御工具，提高网络终端安全防御能力。但是，传统网络安全设备堆叠式的防御手段并不能应对日益多样化、复杂化的网络攻击。相反，

收稿日期：2023-02-28
通信作者：方国强. E-mail：1849336438@qq.com

多异构设备的叠加缺乏统一的视角和关联能力,无法打破数据孤岛,协同防御。因此,系统依然面临着对外服务网站被窜改、敏感数据泄露、恶意程序攻击、分布式拒绝服务攻击等一系列风险。

3　四川气象信息网络安全态势感知系统及关键技术

3.1　系统总体框架

在四川气象信息网络上构建的安全态势感知系统实现了资产管理、安全事件告警、联动响应、通报预警等功能,为网络及时发现问题、处理问题和网络安全决策提供了支撑,平台总体框架如图 2 所示。从态势感知角度,平台分为数据采集及预处理层、数据存储层、态势理解层和态势呈现层。

数据采集及预处理层:包含数据的采集及预处理。平台主要从 3 方面进行数据的采集:(1)在网络中部署探针,用于收集镜像流量,识别出威胁以及审计数据;(2)将防火墙的安全日志和应用控制日志数据上传至平台;(3)全网络范围内部署终端防护工具,将终端安全日志收集并上传至平台。采集到的数据通过清洗、补齐等一系列预处理后,可进行存储。

数据存储层:平台采用 MongDB 作为主要的存储数据库,Elasticsearch 作数据检索库。

态势理解层:该层将 Flink 并行处理和大数据安全分析相结合,对影响网络安全状态的安全事件本身及相关性进行分析,从而预测整个网络的安全态势。

态势呈现层:系统将事件告警、安全态势、通报预警等做成全局性的可视化视图,实现安全事件的实时监控和预警。通过通报预警模块对全网内安全事件进行下发通报,形成安全事件的闭环管理。

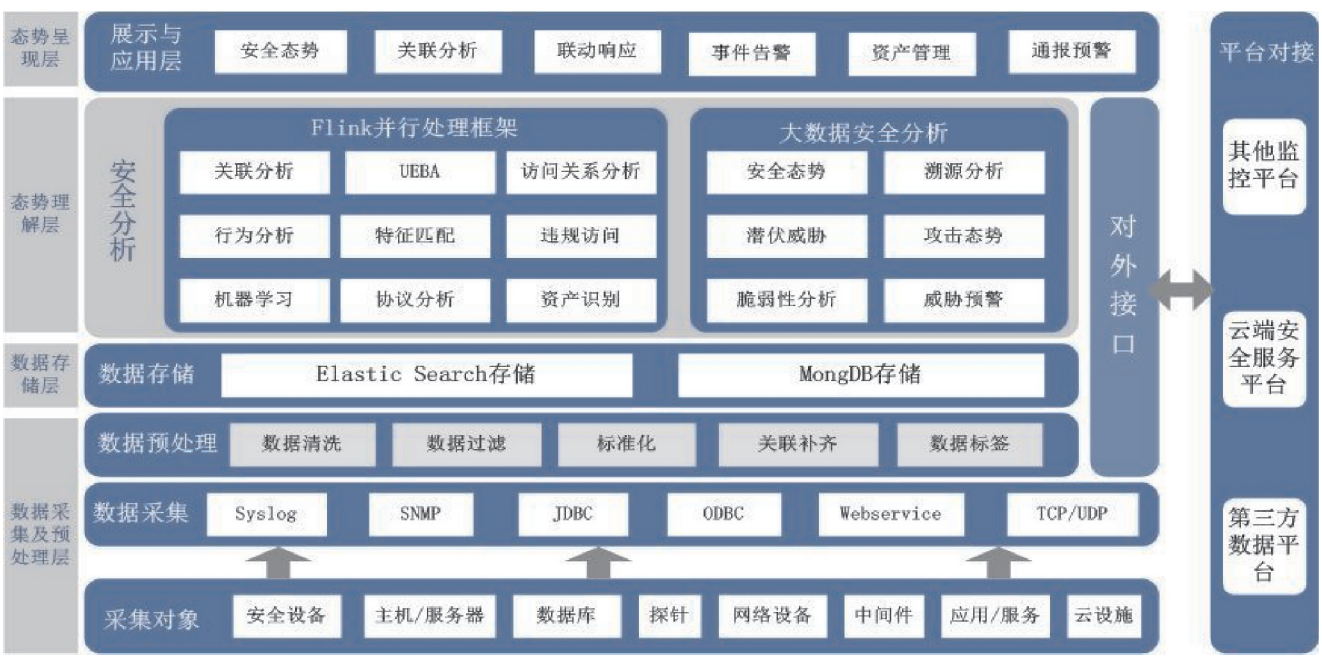

图 2　安全态势感知系统总体框架

3.2　系统关键技术

3.2.1　用户行为画像

用户行为分析[5]主要以用户操作计算机产生的日志数据为输入源,重点考虑用户在做什么、个体行为在时间序列上的关联、用户行为偏好、个体和相似用户群体的差异等方面。通过用户行为画像分析可以预测用户行为,实现异常检测。

系统用户行为画像的主要模块如图 3 所示。系统采集的数据有日志文件直接读取利用、数据库数据利用、日志数据库读取和实时行为数据的采集。对采集的数据进行一些必要预处理,包括数据加载、数据清理、去重、标准化等。对预处理的数据进行聚类、角色标签确定、行为关联分析、频度量化、时间段聚集分析,得到其行为标签。其中,数据分析引擎主要由 Spark 相关组件完成。最后,结合用户角色和日志历史信息勾勒出用户行为画像。根据用户画像分析结果,快速高效地识别用户异常行为。如图 4 所示,系统利用用户行为画像检测出广安气象局某一用户“下载伪造图片文件”这一异常行为。

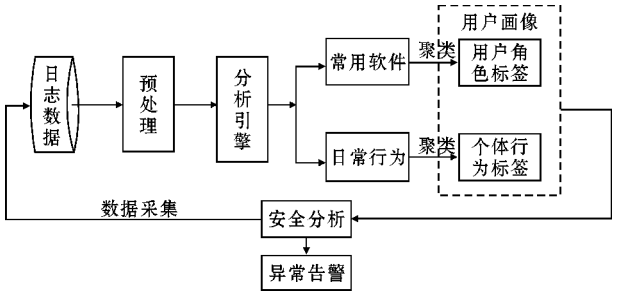

图 3 用户行为画像模块

图 4 用户行为画像分析实例

用户行为画像能有效利用用户历史行为数据,准确实时地判断用户行为异常,提高了系统实时响应能力。

3.2.2 支持向量机

SVM[6]是一种用来解决二分类问题的机器学习算法。系统中,该算法与特征检测相结合应用到邮件安全中,发现伪造邮件、垃圾邮件等威胁。

假设二维空间中数据集 $T=\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$,其中,$x_i\in R^n$,$y_i\in\{+1,-1\}$,$i=1,2,\cdots,N$,$x_i$ 为第 i 个特征向量,y_i 为类标记,等于+1 时为正例,等于−1 时为负例。SVM 算法就是确定一个超平面将两类数据区分开,如图 5 所示。

超平面可通过如下线性方程描述:

$$\omega^T x+b=0 \tag{1}$$

式中:ω 为法向量,决定超平面方向;b 为位移项,决定超平面与原点之间的距离。于是,样本空间中任意点 x 到超平面(ω,b)的距离为

$$r=\frac{|\omega^T+b|}{\|\omega\|} \tag{2}$$

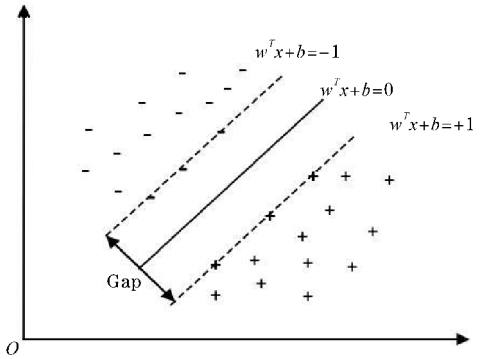

图 5 超平面划分训练样本示意图

假设超平面能将训练样本正确分类,则令对于(x_i,y_i)有

$$\begin{cases}\omega^T x_i+b\geqslant+1, y_i=+1\\ \omega^T x_i+b\leqslant-1, y_i=-1\end{cases} \tag{3}$$

如图 5 所示,距离超平面最近的几个样本点使式(3)等号成立,它们被称为“支持向量”,两个异类支持向量到超平面的距离之和为

$$r=\frac{2}{\|\omega\|} \tag{4}$$

为使 r 最大,ω 则需取最小。所以,求 r 的最大值等价于求下式的最小值:

$$\min\frac{1}{2}\|\omega\|^2 \tag{5}$$

对式(5)添加拉格朗日乘子 α_i,构建拉格朗日函数求解,可得出超平面模型:

$$f(x)=\sum_{i=1}^{n}\alpha_i y_i x_i^T x+b \tag{6}$$

因上述过程满足 KKT 条件,可得:

若 $\alpha_i=0$,则该项不会在式(6)中出现,即该样本不会在最后求解模型参数式子中出现。

若 $\alpha_i>0$,则 $y_i f(x_i)-1=0$,也就是 $y_i f(x_i)=1$。即该样本位于最大间隔边界上,是一个支持向量。

系统利用 SVM 算法检测出的欺诈邮件与垃圾邮件效果如图 6 所示。

序号	时间	描述	日志类型	攻击类型	源IP	源端口	目的IP	目的端口	请求URL	严重等级	数据来源	攻击结果	命中白名单	代
1	2022-11-22 02:01:00	钓鱼邮件	邮件攻击	欺诈邮件	1[redacted].4.11	4442	[redacted].55.59	432	-	高危	办公区[redacted]	尝试	否	-
2	2022-11-22 02:01:00	比特币勒索邮件	邮件攻击	欺诈邮件	[redacted]55.47	23	[redacted].9	234	-	中危	办公区[redacted]	尝试	否	-
3	2022-11-22 02:01:00	垃圾邮件	邮件攻击	垃圾邮件	[redacted]55.55	34	[redacted]	34	-	中危	办公区[redacted]	尝试	否	-
4	2022-11-22 02:01:00	钓鱼邮件	邮件攻击	欺诈邮件	[redacted].55.57	23	[redacted]6	234	-	高危	办公区[redacted]	尝试	否	-
5	2022-11-22 02:01:00	钓鱼邮件	邮件攻击	欺诈邮件	[redacted]0.4.7	2323	[redacted]5.56	122	-	高危	办公区[redacted]	尝试	否	-
6	2022-11-22 02:01:00	钓鱼邮件	邮件攻击	欺诈邮件	[redacted]55.55	34	[redacted]	34	-	高危	办公区[redacted]	尝试	否	-
7	2022-11-22 02:01:00	钓鱼邮件	邮件攻击	欺诈邮件	[redacted]68.4.10	23	[redacted]55.54	234	-	高危	办公区[redacted]	尝试	否	-
8	2022-11-22 02:01:00	钓鱼邮件	邮件攻击	欺诈邮件	[redacted]68.4.4	23	[redacted]5.57	234	-	高危	办公区[redacted]	尝试	否	-
9	2022-11-20 02:00:33	钓鱼邮件	邮件攻击	欺诈邮件	[redacted]68.4.11	4442	[redacted]5.59	432	-	高危	办公区[redacted]	尝试	否	-
10	2022-11-20 02:00:33	比特币勒索邮件	邮件攻击	欺诈邮件	[redacted].55.47	23	[redacted].4.9	234	-	中危	办公区[redacted]	尝试	否	-
11	2022-11-20 02:00:33	垃圾邮件	邮件攻击	垃圾邮件	[redacted].55.55	34	[redacted]	34	-	中危	办公区[redacted]	尝试	否	-

图 6 系统检测邮件类攻击效果图

3.2.3　深度学习

利用卷积神经网络(convolutional neural networks, CNNs)[7]对数亿维的原始特征进行分析、综合,训练出能够进行恶意文件鉴定的模型[8]。

使用灰度图像生成算法 B2M,将大量实验样本即可执行文件转换成相应的灰度图像。算法将任意已知的恶意软件执行文件每8 bit读取为一个不带符号的整型(取值为[0,255]),将固定的行宽设为一个向量,使整个文件最后生成一个二维数组。接着,将训练样本输入 CNN 网络进行模型训练。CNN 结构一般包括卷积层、激活函数层、池化层、全连接层[9],如图 7 所示。

卷积层由若干个卷积单元组成,其通过卷积运算可以提取出图像中的不同特征。假设输入层有 i 个神经元,分别是 $x_1, x_2, \cdots, x_i$。经过卷积层后的输出 X_j 表示为

$$X_j = \sum_i x_i * \omega_{ij} + b_j \tag{7}$$

式中,ω_{ij}表示第 i 个输入神经元进行卷积运算的卷积核,b_j 表示第 j 个输出神经元的偏置值,$*$ 表示卷积运算。

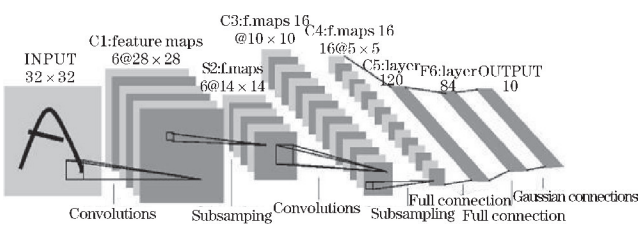

图 7　卷积神经网络结构图

激活函数[10]在前向传播中的作用,是把输入数据的特征通过自身的非线性特性保留并映射出来,传入下一个神经元。在卷积的基础上,一个神经元的输出可以表示为

$$Y_j = \varphi(X_j) \tag{8}$$

池化层(pooling 层)通过池化操作对图像进行降维,并提高图像特征的变换不变特性,通常位于卷积层之后。池化操作是用池化窗口按一定规则在输入特征图中按从左到右,从上到下的顺序进行移动。常用的池化方法有最大池化、均值池化和随机池化 3 类。

全连接层的每一个神经元都与上一层每个神经元连接,把前一层的输出特征都综合起来。所以,该层的权值参数是最多的,其通常在网络的最后几层。

最后,将待检测样本输入训练好的模型进行检测,区分正常文件或恶意文件。系统对恶意文件的检测实例如图 8 所示。

图 8　系统检测恶意文件效果图

4　结束语

针对四川气象信息网络系统存在的安全风险,结合当下热门的网络安全态势感知技术,构建四川气象信息网络安全态势感知系统。系统从安全日志、终端行为、数据流量等多方面入手,进行相关数据的全面收集,通过网络安全态势要素提取、网络安全态势理解、态势风险评估、判断风险等级、态势可视化展现、安全态势预测预警等技术手段,实现了系统对各类安全风险的实时监视,事件预警及通报下发,大幅提升了四川气象信息网络系统监测预警及防御能力。

但是,目前一些正常业务行为也会被系统判定为安全事件,需要人工进一步分析。后续研究可以考虑系统与业务如何进一步融合。此外,对网络安全态势感知技术的研究还处于初级阶段,态势理解的算法还有待深入研究。

参考文献:

[1]　王慧强,赖积保,胡明明,等. 网络安全态势感知关键实现技术研究[J]. 武汉大学学报(信息科学版),2008,33(10):995-998.

[2]　席荣荣,云晓春,金舒原,等. 网络安全态势感知研究综述[J]. 计算机应用,2012,32(1):5.

[3]　石乐义,刘佳,刘祎豪,等. 网络安全态势感知研究综述[J]. 计算机工程与应用,2019,55(24):9.

[4]　贾焰,韩伟红,杨行. 网络安全态势感知研究现状与发展趋势[J]. 广州大学学报(自然科学版),2019,18(3):1-10.

[5]　何雪海,黄明浩,宋飞. 网络安全用户行为画像方案设计[J]. 通信技术,2017,50(4):6.

[6]　丁世飞,齐丙娟,谭红艳. 支持向量机理论与算法研究综述[J]. 电子科技大学学报,2011,40(1):9.

[7] 陈先昌. 基于卷积神经网络的深度学习算法与应用研究[D]. 杭州:浙江工商大学,2014.

[8] 蒋晨,胡玉鹏,司凯,等. 基于图像纹理和卷积神经网络的恶意文件检测方法[J]. 计算机应用,2018,38(10):5.

[9] 常亮,邓小明,周明全,等. 图像理解中的卷积神经网络[J]. 自动化学报,2016,42(9):13.

[10] 田娟,李英祥,李彤岩. 激活函数在卷积神经网络中的对比研究[J]. 计算机系统应用,2018,27(7).

Research on Sichuan Meteorological Information Network Detection and Defense Technology based on Network Security Situation Awareness

TIAN Juan, FANG Guoqiang, HE Xingting, XIE Yinhai

(Sichuan Meteorological Observation Data Center, Chengdu 610072, China)

Abstract: This study presents the development of a Sichuan meteorological information network security situation awareness system based on big data and machine learning technologies. The system provides a comprehensive understanding of the global network state, enhances the monitoring and emergency response capabilities of the meteorological department, and supports informed decision-making. Firstly, this paper introduces the network security situation awareness technology, then expounds onthe current situation and existing risks of meteorological information networks and theirsecurity defense, and then introduces the system and its key technologies from the perspective of practice. Finally, the study summarizes the application of the system and provides future prospects.

Keywords: network security; cyber security situational awareness; meteorological information network; detection and defense